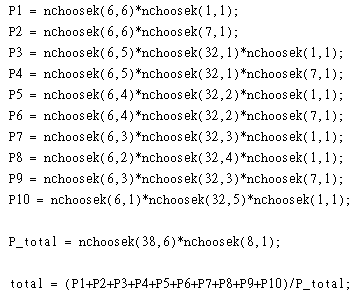
1. According to the winning rules, as shown in Figure 1, the combination of each prize can be written as:

For winning the second set, it is needed that the specific ball should be chosen, which only have 1 combination. But for not choosing the specific ball, the picked ball should be chosen from the others, which have combinations. Same for the first set, if the combination of choosing only m balls from the winning set is times choosing 6-m balls from the losing set .

The total combination of both sets are . By calculating the total probability, we can obtain:



The probability Calculated form MATLAB = 0.1178, near 0.11=

The assumption under this modeling is and iid assumption, meaning that every number has the identical probability to be chosen independently each time.

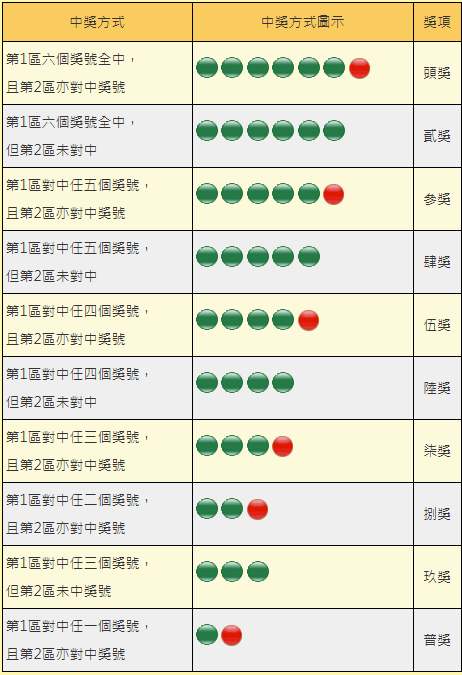


Figure 1



To check if events are independent or mutual exclusive, we first verify if the probability satisfies . If the equation is satisfied, then A and B are independent.

As we can see from the result, only event A and B are independent, both A,C and B,C are dependent.

For all the union of each two event are not empty set, neither of the two combination events are mutual exclusive.

1. From Figure 1, with the right pick from second-set number, we can have the chances to win prize 1, 3, 5, 7, 8, and 10. The condition probability is based on winning the second-set, which has a probability .

The condition probabilities of winning each prize are listed below:

1. (a) For data of first section from the last 50, 100, and 500 runs, the histogram of the number distribution is shown as Figure 2 to Figure 4.

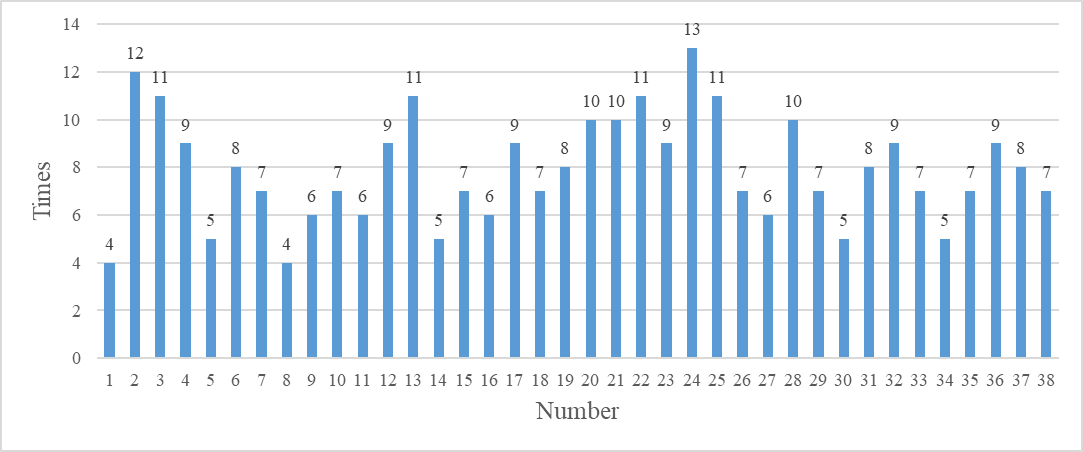


Figure 2 histogram of 50 runs, first section

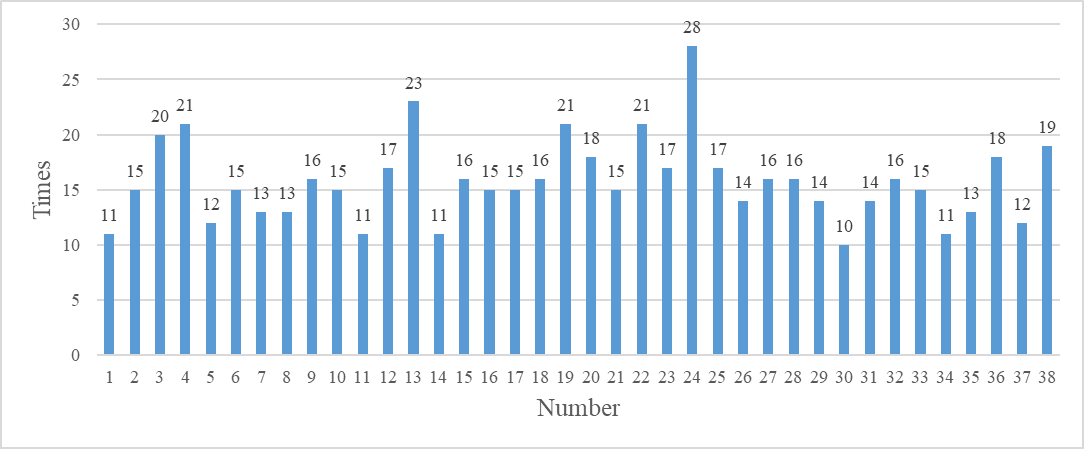


Figure 3 histogram of 100 runs, first section

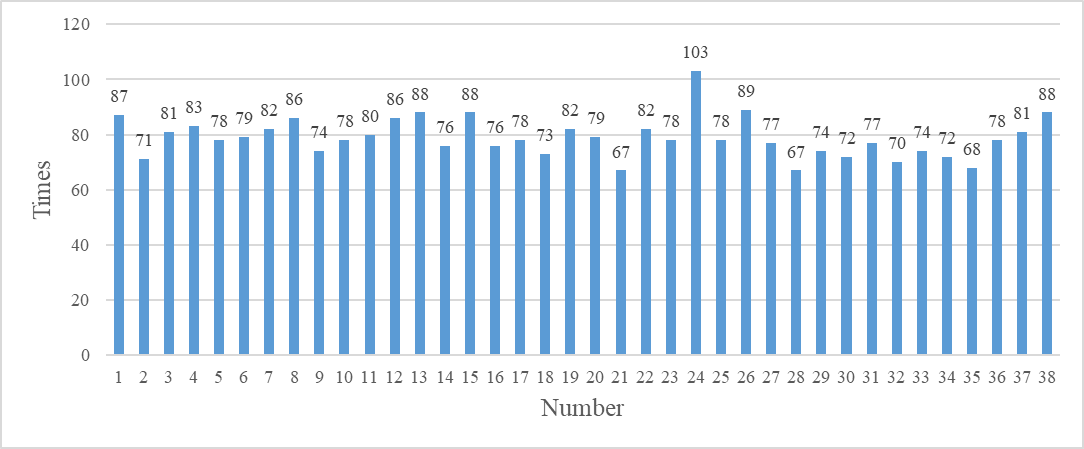


Figure 4 histogram of 500 runs, first section

The statistic identity of the data is shown in the table below:

Table 1 data identity for first section

|  |  |  |  |
| --- | --- | --- | --- |
|  | 50 | 100 | 500 |
| avg | 7.894737 | 15.78947 | 78.94737 |
| medium | 7.5 | 15 | 78 |
| max | 13 | 28 | 103 |
| min | 4 | 10 | 67 |
| range | 9 | 18 | 36 |
| variance | 5.069701 | 13.95448 | 52.10526 |

(b) For data of second section from the last 50, 100, and 500 runs, the histogram of the number distribution is shown as Figure 5 to Figure 7.

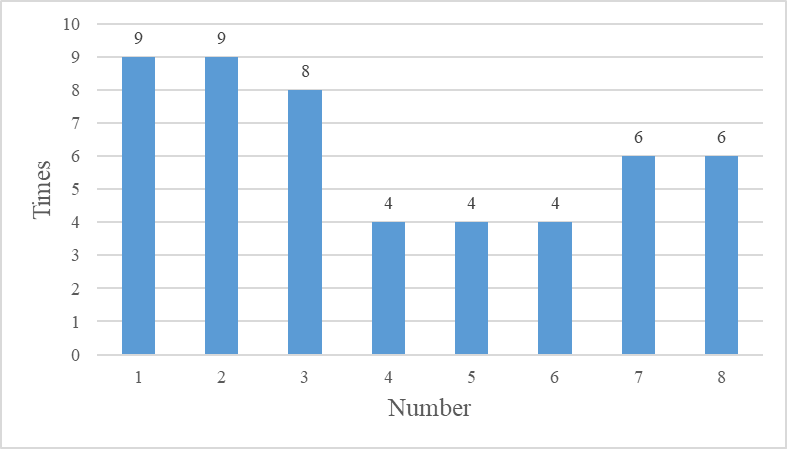


Figure 5 histogram of 50 runs, first section

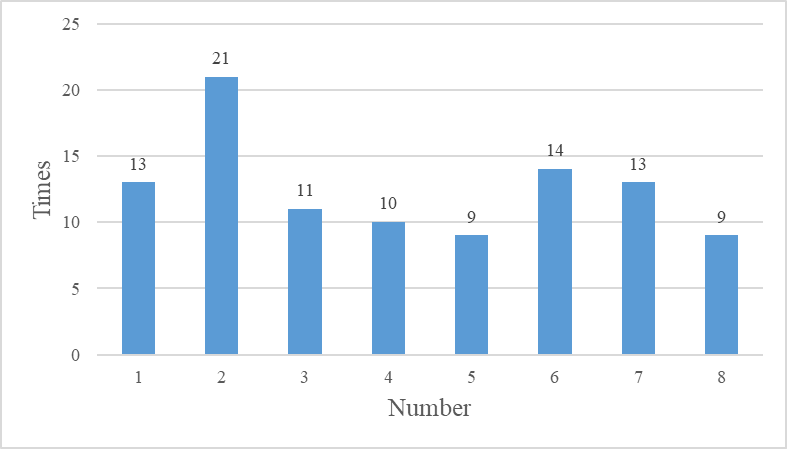


Figure 6 histogram of 100 runs, first section

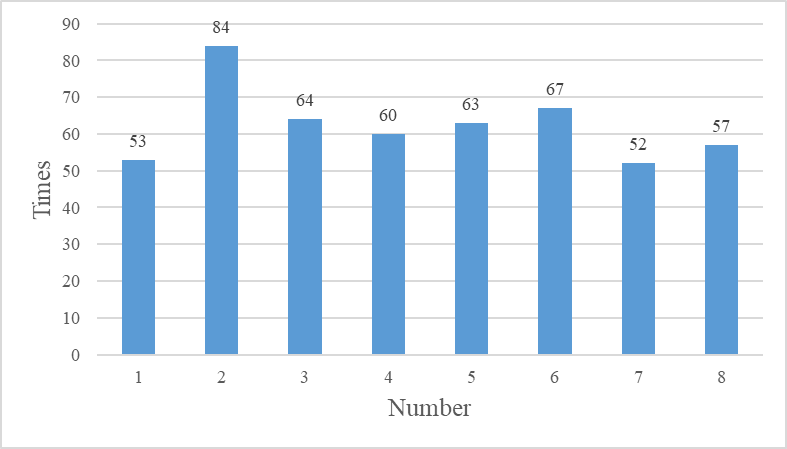


Figure 7 histogram of 500 runs, first section

The statistic identity of the data is shown in the table below:

Table 2 data identity for second section

|  |  |  |  |
| --- | --- | --- | --- |
|  | 50 | 100 | 500 |
| avg | 6.25 | 12.5 | 62.5 |
| medium | 6 | 12 | 61.5 |
| max | 9 | 21 | 84 |
| min | 4 | 9 | 52 |
| range | 5 | 12 | 32 |
| variance | 4.785714286 | 15.42857143 | 103.1428571 |